Visual Search Distributions

Robert E. Karlsen
U.S. Army – Tank Automotive Research, Development and Engineering Center
Warren, MI 48397-5000

ABSTRACT

In this paper, we explore the distribution of search times in narrow-field-of-view visual search tasks. In these types of experiments, it is typical to perform tests with multiple observers examining multiple scenes. The results are aggregated over all observers to find estimates of target conspicuity in each scene. Methods such as signal detection theory have been developed to take into account variations in observers. However, many of these methodologies are technically only applicable if certain strict conditions are met. If one instead is interested in observer response time distributions, one generally averages over all the images as well, in order to obtain a sufficient number of data points. The question we address in this paper is whether one can say anything meaningful about the search time distribution for individual images, based on the aggregated search time distribution over all the images.

INTRODUCTION

The modeling approach remains within the general framework of the Night Vision Laboratory (NVL) search modeling methodology. We investigate alternative model formulation details in the light of experimental data from a large-scale search and target acquisition experiment. The scope of this search experiment was limited to narrow-field-of-view search, roughly corresponding to the macular visual field (approximately 10 degrees of visual angle). As with the NVL methodology, the focus is on modeling aggregate performance over an ensemble of observers.

The NVL search model assumes a negative exponential distribution of response time. Our data show that the negative exponential distribution overestimates the probability of fast detections, and that search time more closely follows a lognormal distribution. While a preliminary analysis supports this conclusion[1], we have not yet completed a more detailed statistical analysis. The data show that probability of detection in search is highly correlated with probability of detection given a cue to the target location, but slightly lower. This indicates that in some cases, observers who could detect the target if given a cue failed to find the target during search.

UNCUED SEARCH AND CUED DISCRIMINATION EXPERIMENTS

The results from two separate discrimination experiments performed by Turing Associates are described here. The experiments have been described in detail elsewhere[2-4] and therefore, are only briefly summarized in this paper. In the search experiments, search was terminated at the first reported detection. False positives were reported for scenes with targets, so it is possible that false alarms preempted target detection in a fraction of the cases.

The experiments were based on a set of 1150 images of military vehicles in natural terrain. The images were constructed from a base set of 44 photographs taken at NVL's 1995 DISSTAF (Distributed Interactive Simulation Search and Target Acquisition Fidelity) field trials at Fort Hunter Liggett in California. There were nine types of military vehicles including tanks, armored personnel carriers, and trucks of U.S. and Soviet manufacture. The vehicles were in tactically appropriate positions exploiting cover and concealment, at ranges from 500 to 5,000 meters. Photographs were taken with a 10x lens on 35mm film. The images were professionally digitized to 6144 by 4045 pixels.

The images were copied and digitally modified to increase the range of target detectability, and to increase the variety of visual conditions. Modified versions were created by attenuating luminance, contrast and color. The images were also modified by digital resampling at two levels to simulate the effect of different viewing ranges. Additional modifications

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
16 SEP 2003	**Journal Article**	**16-09-2003 to 06-09-2003**

4. TITLE AND SUBTITLE **Visual Search Distributions**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) **Robert Karlsen**	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army TARDEC ,6501 E.11 Mile Rd,Warren,MI,48397-5000**	8. PERFORMING ORGANIZATION REPORT NUMBER **#14250**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **U.S. Army TARDEC, 6501 E.11 Mile Rd, Warren, MI, 48397-5000**	10. SPONSOR/MONITOR'S ACRONYM(S) **TARDEC** 11. SPONSOR/MONITOR'S REPORT NUMBER(S) **#14250**

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

Submitted to Ground Target Modeling Conference 2003

14. ABSTRACT

In this paper, we explore the distribution of search times in narrow-field-of-view visual search tasks. In these types of experiments, it is typical to perform tests with multiple observers examining multiple scenes. The results are aggregated over all observers to find estimates of target conspicuity in each scene. Methods such as signal detection theory have been developed to take into account variations in observers. However, many of these methodologies are technically only applicable if certain strict conditions arc met. If one instead is interested in observer response time distributions, one generally averages over all the images as well, in order to obtain a sufficient number of data points. The question we address in this paper is whether one can say anything meaningful about the search time distribution for individual images, based on the aggregated search time distribution over all the images.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**7**	

were made by attenuating the signature of the target. In the final image set, the target was replaced by appropriate background in approximately one-third of the images. The images were then cropped to 1080 by 720 pixels so that the target locations were approximately uniformly distributed over the image. This was done to prevent anticipation of the target location and to disrupt learning effects. Half of the images were flipped horizontally to also minimize any learning effects.

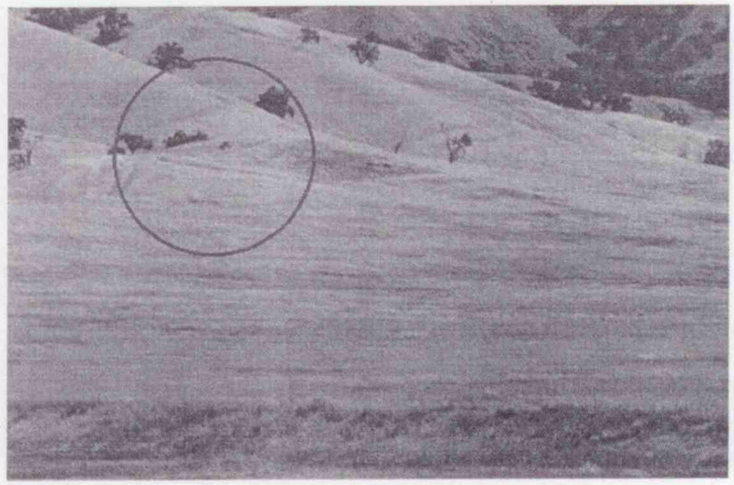


Figure 1: Image from cued detection experiment.

The experiments used 30 college-age subjects. Different subjects were used for each of the experiments. Subjects were tested individually with images displayed on a computer monitor. Prior to testing, the subjects were thoroughly trained by allowing them to study a representative set of images of the targets in background and close-up images of the targets from a variety of angles. The training images were not used in the testing. In the experiments the subjects were seated 60 inches from the monitor, resulting in a resolution of 110 pixels per degree of visual angle. The experiments were self-paced and timed out at 60 seconds.

In the cued discrimination test, the target, if present, was at the center of a red circle 3 degrees in diameter, as depicted in Fig. 1. For no-target or background-only images, the cueing circle was centered on a typical vehicle location. The subjects were given a four-level confidence scale: (0) no vehicle present, (1) unsure whether a vehicle was present, (2) confident a vehicle was present, and (3) certain a vehicle was present.

In the uncued search test, the subjects were instructed to move the cursor and click on the target if present, or to click anywhere if no target was detected. After clicking, the subjects were presented with the same 4-choice menu as used in the cued discrimination test. Response time, click location and menu choice were recorded.

Click location was used to distinguish correct detections from false alarms. When the click was within 55 pixels of the target, the response was considered a correct detection. When the click was more than this from the target, it was considered a false alarm. This is a simplistic method that can result in misclassification of responses. A more accurate method would be to use the dimensions of the target in determining click precision. This will be implemented in future analyses.

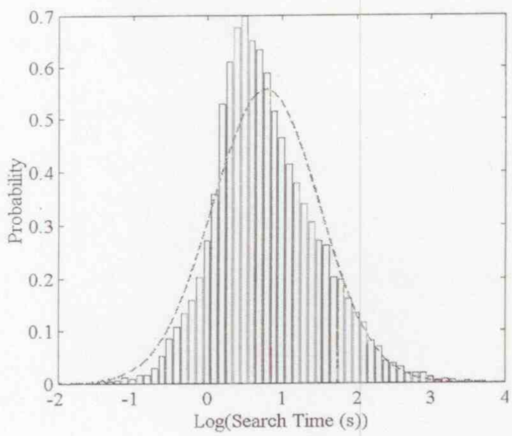
Seven of the subjects either did not comply with the instructions or had equipment malfunctions and their results were discarded. This left 23 subjects in the population sample.

DISTRIBUTION OF RESPONSE TIME

The Night Vision Laboratory's (NVL) search and static performance models are the standard tools to predict search and target acquisition performance. The models are simple and widely used. They are useful as a reference and starting point to explain the advances in search modeling.

Search is conceptualized as consisting of two alternating activities: scanning the scene to find possible targets, and inspecting them to decide whether or not to report a target detection[5]. The NVL search model predicts the probability of detection in less than time T as the product of the limiting probability of detection, P_{inf}, and the distribution of response time[6,7]. The NVL search model assumes that response time has a negative exponential cumulative distribution function (CDF):

$$P_d(t < T) = P_{\text{inf}}(1 - e^{-T/\tau}), \tag{1}$$



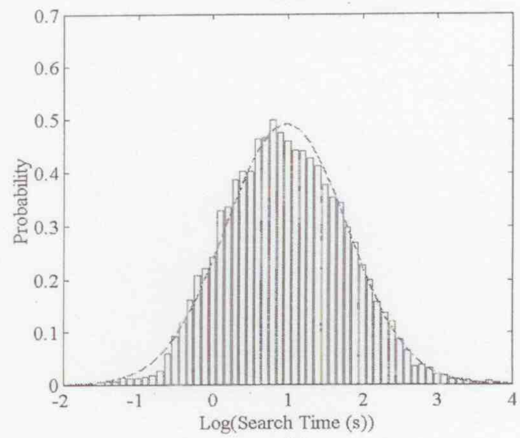


Figure 2: Search time distributions for (a) images with targets and for (b) images without targets.

where τ is the mean response time. In static detection assessment, i.e., without search, P_{inf} is commonly interpreted as the fraction of the observer population that can detect the target in unlimited time given a cue to the target location. In search modeling, P_{inf} is interpreted as the fraction of the observer population who will eventually detect the target given unlimited time searching for all possible targets in the scene (so that target detection can not be preempted by a false alarm). The two interpretations are equivalent under the assumption that all the observers capable of detecting the target in the static detection scenario will also find the target and inspect it during search.

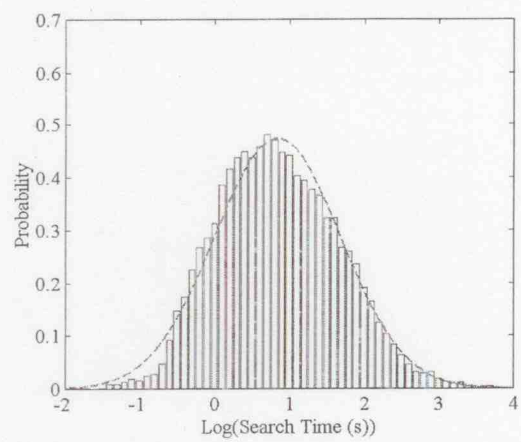


Figure 3: Search time distribution for "no target" response.

However, experimental psychologists have observed that the distribution of search time often resembles a lognormal distribution and have shown that there are a variety of conceptual mechanisms that could produce lognormally distributed response times[8,9]. The general argument is that in visual search of complex scenes, the rate of response is initially low, before the observer has had time to understand the scene. As observers begin to search for targets, they also build up an understanding of the scene that enables them to search more productively by excluding irrelevant regions and focusing on likely areas. Search and scene parsing proceed from coarse to fine, a strategy that leads to high payoff early in search and rapid detection of larger and higher signature targets. These behaviors lead to a rapid increase in the rate of response, followed by a decrease once the scene has been parsed and search devolves into serial search for low-signature targets. As the observers lower their thresholds to look for lower signature targets, the rate of response continues to decrease.

The lognormal distribution accounts for an early "ramp-up" stage of search, a period of high productivity, followed by a gradual decline in productivity. A lognormal response time distribution is equivalent to a normal distribution of the log of the response time:

$$P(t < T) = \frac{1}{\sqrt{2\pi}\sigma}\int_0^T \frac{1}{t} e^{-(\log t-\mu)^2/2\sigma^2}\,dt = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\log T} e^{-(z-\mu)^2/2\sigma^2}\,dz\,, \tag{2}$$

where μ is the mean log response time and σ is the standard deviation of the log response time. Figures 2 show the empirical probability density function compiled from all subjects for images with targets and for images without targets. It is seen that for those images with targets, the distribution of the logarithm of search time was decidedly non-normal, while for those

images without targets, the distribution appears closer to normal. Figure 3 contains response data where the subject said that there was no target, and the data includes images both with a target and without a target. This means essentially that the subject scanned the entire image and was not able to find an object sufficiently target-like. It was conjectured that the subject response distribution should be similar in these cases, regardless of the actual presence of a target or not. The distribution in this case appears close to normal as well.

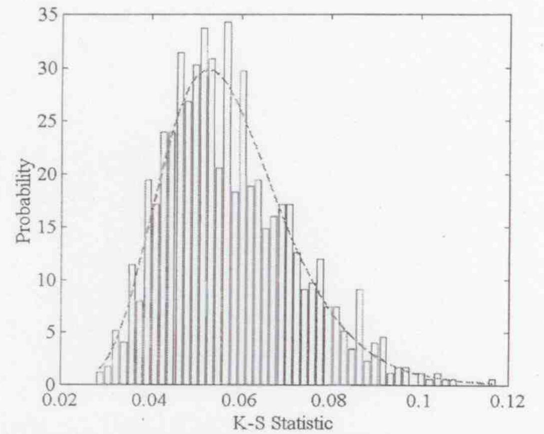


Figure 4: Distribution of K-S statistic for 100 samples drawn from a normal distribution.

KOLMOGOROV-SMIRNOV STATISTIC

In many cases, instead of the aggregate distribution, one wants the distribution of search times for each image. The question we address then is whether one can deduce any information about the individual distributions from the aggregate distribution. And whether the aggregate distribution can provide any more information than by analyzing the individual image distributions by themselves.

To begin, it is necessary to go beyond visually comparing the experimental histogram to the hypothetical distribution. One needs a metric to compare the distributions systematically and objectively. A metric is also important when the sample size is too small to support plotting a histogram. There are actually many methods to do this. We chose to use the well-known Kolmogorov-Smirnov (K-S) statistic[10] and a variant due to Kuiper[10]. The K-S statistic is straightforward to compute; it is simply the largest absolute deviation between the cumulative distributions,

$$D_{KS} = \max\left|F_n(x) - F(x)\right|, \tag{3}$$

where $F(x)$ is the conjectured cumulative distribution and $F_n(x)$ is the sampled cumulative distribution. The latter is simply the fraction of data points less than x and is easily computed by sorting the data. The K-S statistic has the property that it is scale invariant and it can be applied to any distribution, although it may not be accurate for some pathological cases. However, the sensitivity of the K-S statistic is dependent on x and in fact is less sensitive in the tails of the distribution than at the peak. This has led to an assortment of variations of the K-S statistic to alleviate this shortcoming. One variant, due to Kuiper[10] (KKS), is also relatively simple to implement. It is simply the difference between the maximum difference and the minimum difference or,

$$D_{KKS} = \max\left(F_n(x) - F(x)\right) + \max\left(F(x) - F_n(x)\right). \tag{4}$$

The sensitivity of the KKS statistic is independent of x.

Since we are interested in comparing distributions to a normal distribution, it is necessary to determine what the statistics of the statistic is, as a function of the number of samples, for a normal distribution. We do so by generating a large number of normal distributions of various lengths and then compute the KS and KKS statistics for each of them. A histogram analysis reveals that both statistics have a lognormal shape, as seen in Fig. 4 for the case of 100 samples for the KS statistic. And indeed, if one computes the KS and KKS statistic, one finds that the statistic of the log(statistic) is consistent with a normal distribution. Computing the mean and standard deviation of the log(statistic) then gives a measure of the distribution of the statistic. If we plot the mean and the three standard deviation (3-sigma) limit of the log(statistic) versus the number of samples, we find that a power law provides a reasonable fit to the data, as seen in Fig. 5

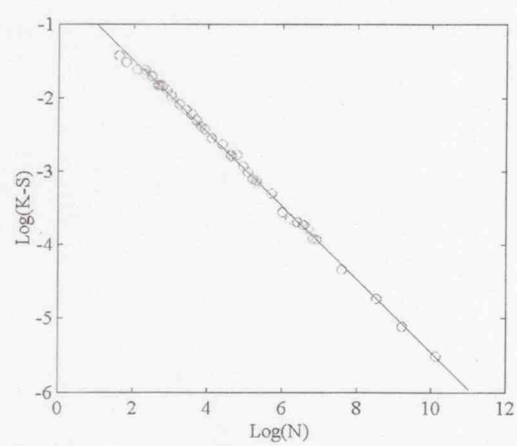


Figure 5: Regression plot for the three standard deviation limit of the K-S statistic of a normal distribution versus the number of samples.

for the 3-sigma limit. Based on this fit, we can compute the mean and three standard deviation limit for a normal distribution for any number of samples. Recall that the 3-sigma limit for a one-tailed distribution is equivalent to a probability of 0.0013 that the distribution in question is drawn from a normal distribution.

Computing the K-S statistic for the distributions in Figs. 2 and 3, we find that KS=0.0535 for Fig. 2(a), KS=0.0171 for Fig. 2(b), and KS=0.0219 for Fig. 3. The 3-sigma limits for normal distributions of these particular sample sizes are 0.0096, 0.0145, and 0.0106, respectively. Therefore we conclude each of the three is unlikely to have been drawn from a normal distribution, with the no-target data in Fig. 2(b) the closest to normality. Note that we have found that the Kuiper variant of the K-S statistic gave essentially the same result for our data as did the K-S statistic and therefore, in the rest of the paper, we will only report K-S results.

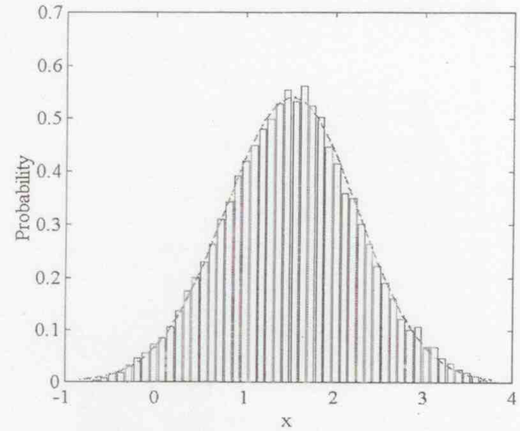


Figure 6: Aggregate distribution formed from uniform distributions whose position and width were drawn from uniform distributions.

AGGREGATE DISTRIBUTIONS

To address the question of whether the aggregate distribution has any relationship to the underlying individual distributions, we created functions to produce three types of sample distributions, a normal distribution, a uniform distribution and a distribution based on a polynomial function. We used one of these three for the underlying distribution and one to model the distribution of the mean and the width of the underlying distributions. We generated a list of means and widths, based on the chosen distribution, and then drew numbers from these lists to create a large number of distributions that were aggregated into a single distribution. This aggregate distribution is than analyzed to determine its statistics. Two combinations demonstrate the point we are trying to make. Figure 6 shows the aggregate distribution formed from a group of uniform distributions, where the mean and the width were also taken from uniform distributions. The K-S statistic for this data is 0.0053, while the 3-sigma limit is 0.0083, indicating that the resulting aggregate distribution is likely to have come from a normal distribution. It should be pointed out that by adjusting some of the parameters slightly for the mean or width distributions, the aggregate distribution in this case quickly becomes non-normal. On the other hand, Figure 7 shows the aggregate distribution where the underlying distributions are normal and the mean and width distributions are also drawn from normal distributions. Here the K-S statistic is 0.0284 with the same 3-sigma limit of 0.0083, indicating that the resulting aggregate distribution is unlikely to have come from a normal distribution. These two examples and other runs that we have made show that the shape of the aggregate distribution has no decisive relationship to the shape of the underlying distributions or to the distribution of the mean and width parameters. However, we have not performed a sensitivity analysis to determine what the probabilities are for a given underlying distribution, given a particular shape of the aggregate distribution. It may turn out that the result is particularly sensitive to the specific set of parameters used to generate the individual distributions.

Based on the previous analysis, we applied the K-S analysis to the observer data for each image. The histograms for the mean and standard deviation showed that neither are likely to have been drawn from a normal distribution. The K-S statistic for these distributions supports this conclusion. The distribution of the K-S statistic on the other hand is consistent with a lognormal distribution, as seen in Fig. 8, much like the simulated normal results in Fig. 4. The KS value for the images with targets in Fig. 8(a) is 0.0247 with a 3-sigma limit of 0.0429, while the KS value for the images without targets in Fig. 8(b) is 0.0287 with a 3-sigma limit of 0.0649. Additionally, note that the vast majority of the KS values for the individual image data are consistent with being drawn from a normal distribution with 23 samples, where

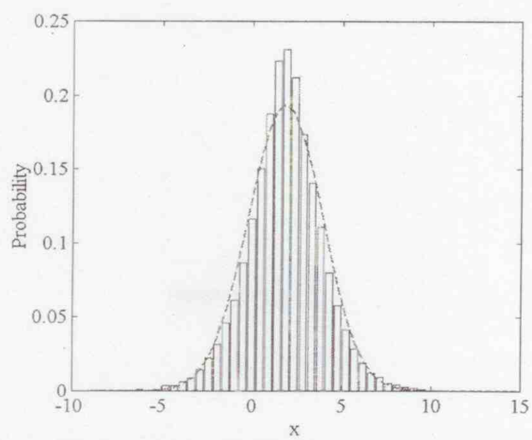


Figure 7: Aggregate distribution formed from normal distributions whose mean and standard deviation were drawn from normal distributions.

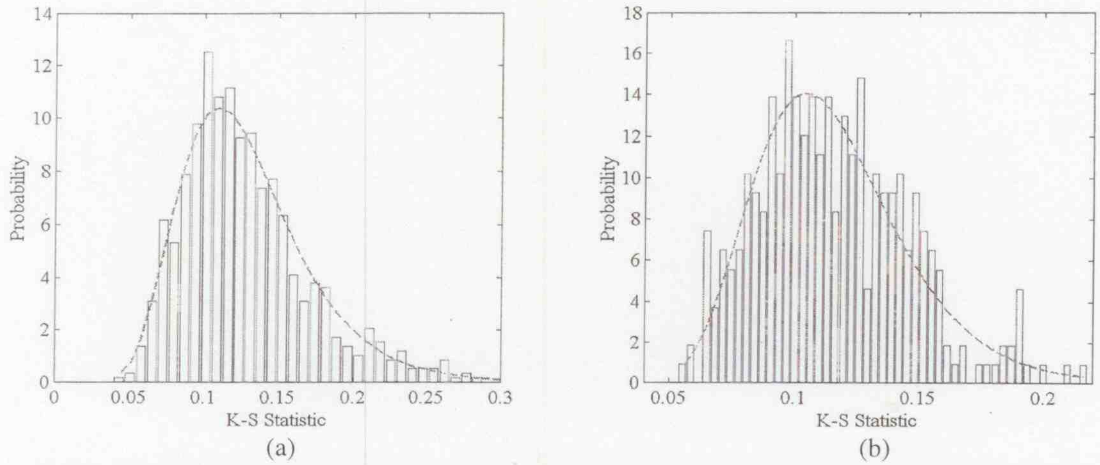


Figure 8: Distribution of the K-S statistic for (a) images with targets and for (b) images without targets.

the mean is 0.107 and the 3-sigma limit is 0.235. So for this observer data, while the aggregate distributions in Fig. 2 are non-normal, the majority of the underlying individual distributions are consistent with a normal distribution.

CONCLUSION

In this paper we have shown that it incorrect to assume that the form of an aggregate distribution correlates with the form of the underlying individual distributions. We demonstrated this on two simulated data sets, both showing that aggregations of non-normal distributions can result in a normal distribution and that aggregations of normal distributions can result in a non-normal distribution.

We also showed an analysis of a set of search times for a narrow field of view search test that supported the hypothesis that search times follow a lognormal distribution. We used the Kolmogorov-Smirnov statistic[10] to test for normality, as well as the variant due to Kuiper[10], although the latter did not contradict the former in any of the data that we analyzed.

Further work will consist of a more careful analysis of the tests for normality, a parametric sensitivity analysis on the effect of different distributions, and an analysis into a possible correlation between normality and certain image characteristics.

REFERENCES

1) G. Witus, R.E. Karlsen, G.R. Gerhart and D.J. Gorsich, "Advances in modeling visual search and target discrimination performance," *Proc. 13th Ground Target Modeling and Validation Conf.*, ed. W. Reynolds, Houghton, MI, 247-56 (1999).

2) R.D. Ellis and G. Witus, "A robust data set for visual detection model calibration and validation," *Proc. 10th Ground Target Modeling and Validation Conf.*, ed. W. Reynolds, Houghton, MI, 230-32 (1999).

3) G. Witus and D. Ellis, "Perception testing for development of computer models of ground vehicle visual discrimination performance," SPIE Proc. **4029**, *Targets and Backgrounds VI*, eds. W.R. Watkins, D. Clement and W.R. Reynolds, 2-9 (2000).

4) G. Witus and R.D. Ellis, "Visual search and cued detection performance," *SPIE Proc.* **4370**, *Targets and Backgrounds VII*, eds. W.R. Watkins, D. Clement and W.R. Reynolds, 1-9 (2001).

5) B.L. O'Kane, "Validation of Prediction Models for Target Acquisition with Electro-Optical Sensors," *Vision Models for Target Detection and Recognition*, ed. E. Peli, World Scientific, Salem, MA, 192-218 (1995).

6) J. Maaz, "Acquire model: variability in N_{50} analysis," *Proc. 9th Ground Target Modeling and Validation Conference*, ed. W. Reynolds, Houghton, MI, 36-42 (1998).

7) M. Friedman, "Recommended technique for calculating τ, P_{inf} and their uncertainites in search with no external time limit," *Proc. 11*

8) R. Ulrich and J. Miller, "Information processing models generating lognormally distributed response times," *Journal of Mathematical Psychology* **37**, 513-25 (1993).

9) G.J.P. Van Breukelen, "Parallel information processing models compatible with lognormally distributed response times," *Journal of Mathematical Psychology* **39**, 396-99 (1995).

10) See e.g., W.H. Press et al, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press (1992).